

Network Working Group	S. Bryant	
Internet-Draft	M. Shand	
Intended status: Informational	Cisco Systems	
Expires: May 4, 2009	October 31, 2008	

[TOC](#)

Applicability of Loop-free Convergence draft-bryant-shand-lf-applicability-06

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/lid-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

This Internet-Draft will expire on May 4, 2009.

Abstract

This draft describes the applicability of loop free convergence technologies to a number of network applications.

Table of Contents

- [1.](#) Introduction
- [2.](#) Applicability
 - [2.1.](#) Component Failure
 - [2.2.](#) Component Repair
 - [2.3.](#) Managing withdrawal of a component
 - [2.4.](#) Managing Insertion of a Component
 - [2.5.](#) Managing Change of a Link Cost
 - [2.6.](#) External Cost Change

2.7.	MPLS Applicability
2.8.	Routing Vector and Path Vector Convergence
3.	IANA considerations
4.	Security Considerations
5.	Informative References
§	Authors' Addresses
§	Intellectual Property and Copyright Statements

1. Introduction

[TOC](#)

When there is a change to the network topology (due to the failure or restoration of a link or router, or as a result of management action) the routers need to converge on a common view of the new topology, and the paths to be used for forwarding traffic to each destination. During this process, referred to as a routing transition, packet delivery between certain source/destination pairs may be disrupted. This occurs due to the time it takes for the topology change to be propagated around the network together with the time it takes each individual router to determine and then update the forwarding information base (FIB) for the affected destinations. During this transition, packets may be lost due to the continuing attempts to use the failed component, and/or due to forwarding loops. Forwarding loops arise due to the inconsistent FIBs that occur as a result of the difference in time taken by routers to execute the transition process. This is a problem that occurs in both IP networks and MPLS networks that use LDP [\[RFC5036\]](#) (Andersson, L., Minei, I., and B. Thomas, "LDP Specification," October 2007.) as the label switched path (LSP) signaling protocol.

The service failures caused by routing transitions are largely hidden by higher-level protocols that retransmit the lost data. However new Internet services are emerging which are more sensitive to the packet disruption that occurs during a transition. To make the transition transparent to their users, these services require a short routing transition. Ideally, routing transitions would be completed in zero time with no packet loss.

Regardless of how optimally the mechanisms involved have been designed and implemented, it is inevitable that a routing transition will take some minimum interval that is greater than zero. This has led to the development of a TE fast-reroute mechanism for MPLS [\[RFC4090\]](#) (Pan, P., Swallow, G., and A. Atlas, "Fast Reroute Extensions to RSVP-TE for LSP Tunnels," May 2005.). Alternative mechanisms that might be deployed in an MPLS network and mechanisms that may be used in an IP network are work in progress in the IETF [\[I-D.ietf-rtgwg-ipfrr-framework\]](#) (Shand, M. and S. Bryant, "IP Fast Reroute Framework," October 2009.). Any repair mechanism may however be disrupted by the formation of micro-

loops during the period between the time when the failure is announced, and the time when all FIBs have been updated to reflect the new topology.

This disruptive effect of micro-loops led the IP fast re-route designers to develop mechanisms to control the re-convergence of networks in order to prevent disruption of the repair and collateral damage to other traffic in the network [\[I-D.bryant-shand-lf-conv-frmwk\]](#) (Bryant, S. and M. Shand, "A Framework for Loop-free Convergence," October 2006.), [\[I-D.ietf-rtgwg-microloop-analysis\]](#) (Zinin, A., "Analysis and Minimization of Microloops in Link-state Routing Protocols," October 2005.).

The purpose of this note is to draw the attention of the IETF community to the more general nature of the micro-looping problem, and the wider applicability of loop-free convergence technology.

2. Applicability

[TOC](#)

Loop free convergence strategies are applicable to any problem in which inconsistency in the FIB causes the formation of micro-loops.

For example, the convergence of a network following:

- 1) Component failure.
- 2) Component repair.
- 3) Managing withdrawal of a component.
- 4) Managing insertion of a component.
- 5) Management change of link cost (either positive or negative).
- 6) External cost change, for example change of external gateway as a result of a BGP change.
- 7) A shared risk link group (SRLG) failure.

In each case, a component may be a link or a router.

2.1. Component Failure

[TOC](#)

When fast-reroute is used to provide the temporary repair of a failed component, the use of a loop-free convergence mechanism enables the re-convergence of the network to be performed without additional packet loss caused by starvation or micro-looping.

The need for loop-free convergence was first appreciated during the design of IP fast reroute. However the mechanism is also applicable to the case where an MPLS-TE tunnel is used to provide a link or node repair within an MPLS network where LDP is used to distribute labels. Except in special circumstances, controlled convergence in the presence of component failure should only be used when a temporary repair is available. This is because controlled convergence is always slower than uncontrolled (traditional) convergence, and would result in an extended period of traffic lost as a result of the failure if there were no other means of delivering the traffic.

2.2. Component Repair

[TOC](#)

Micro-loops may form when a component is (re)introduced into a network. All of the known loop-free convergence methods are capable of avoiding such micro-loops. It is not necessary to employ any repair mechanism to take advantage of this facility, because the new component may be used to provide connectivity before its presence is made known to the rest of the network.

2.3. Managing withdrawal of a component

[TOC](#)

From the perspective of the routing protocol, management withdrawal of a component is indistinguishable from an unexpected component failure, and will be subject to the same micro-loops. The network will therefore benefit from the use of a micro-loop prevention mechanism.

Unlike the failure case, the component being withdrawn may be used to forward packets during the transition, and therefore no repair mechanism is needed.

Unlike the case of component failure or repair, management withdrawal of a component is normally not time critical. Consideration may therefore be given to the use of the incremental cost change loop-free convergence mechanism. This mechanism was discarded as a candidate in the case of fast re-route because of its slow time to converge, however it is a mechanism that is backwards compatible with existing routers and may therefore be of use in this application. Note that unlike any of the other mechanism described here, this technique can be used without modification to ANY router in the network.

[TOC](#)

2.4. Managing Insertion of a Component

From the perspective of the routing protocol, management insertion of a component is indistinguishable from component repair, and will be subject to the same micro-loops. The network will therefore benefit from the use of a micro-loop prevention mechanism. No repair mechanism is needed and it is not normally time critical.

2.5. Managing Change of a Link Cost

[TOC](#)

Component failure and component repair are extreme examples of cost change. Micro-loops may also form when a link cost is changed (in either direction) during the process of network re-configuration. The use of a loop-free convergence technique prevents the formation of micro-loops during this otherwise benign process. No repair mechanism is needed in this case, because the link is still available for use.

2.6. External Cost Change

[TOC](#)

An external cost change can result in a change to the preferred external route to a destination. Micro-loops may form during the process of switching from the old border router to the new one. The loop-free control of this change will prevent the loss of packets during this network transition.

2.7. MPLS Applicability

[TOC](#)

Where the network is an MPLS enabled network using the LDP protocol to learn labels, and fast re-route is provided through the use of single hop MPLS-TE tunnels protected by MPLS-TE fast reroute, micro loops may form during convergence. Loop free convergence is therefore applicable to this network configuration.

2.8. Routing Vector and Path Vector Convergence

[TOC](#)

The work to date on controlled convergence has focused on link state IGP's. The ability to control the convergence of routing vector and path vector routing protocols would also be useful tools in the management of the Internet.

Routing vector convergence may be controlled by incrementally changing the link cost one unit at a time and waiting for the network to converge. In link state routing protocols employing wide metrics such a solution would normally be considered as too slow to deploy, although recent work on optimizing the number of increments has significantly improved the convergence time. In the case of distance vector routing protocols the much smaller maximum metric makes this more tractable, provided that is, the per metric change convergence time is considered acceptable.

Similarly with path vector routing protocols the path length can be incrementally padded. Since practical path vector routing protocols which use path length as an input to the routing decision are equivalent to using the hop count as a metric (i.e. the maximum per hop metric is one ,or in special cases a very small number) the number of increments needed is limited to the length of the path around the failure. This property may make this a tractable approach.

[\[I-D.bryant-shand-lf-conv-frmwk\]](#) (Bryant, S. and M. Shand, "A Framework for Loop-free Convergence," October 2006.) (Section 5.1), although the per change convergence time may still be a significant concern.

3. IANA considerations

[TOC](#)

There are no IANA considerations that arise from this draft.

4. Security Considerations

[TOC](#)

All micro-loop control mechanisms raise significant security issues which must be addressed in their detailed technical description.

5. Informative References

[TOC](#)

[I-D.bryant-shand-lf-conv-frmwk]	Bryant, S. and M. Shand, " A Framework for Loop-free Convergence ," draft-bryant-shand-lf-conv-frmwk-03 (work in progress), October 2006 (TXT).
[I-D.ietf-rtgwg-ipfrr-framework]	Shand, M. and S. Bryant, " IP Fast Reroute Framework ," draft-ietf-rtgwg-ipfrr-framework-13 (work in progress), October 2009 (TXT).
[I-D.ietf-rtgwg-microloop-analysis]	Zinin, A., " Analysis and Minimization of Microloops in Link-state Routing Protocols ," draft-ietf-rtgwg-microloop-analysis-01 (work in progress), October 2005 (TXT).
[RFC4090]	

	Pan, P., Swallow, G., and A. Atlas, " Fast Reroute Extensions to RSVP-TE for LSP Tunnels ," RFC 4090, May 2005 (TXT).
[RFC5036]	Andersson, L., Minei, I., and B. Thomas, " LDP Specification ," RFC 5036, October 2007 (TXT).

Authors' Addresses

[TOC](#)

	Stewart Bryant
	Cisco Systems
	250, Longwater, Green Park,
	Reading RG2 6GB, UK
	UK
Email:	stbryant@cisco.com
	Mike Shand
	Cisco Systems
	250, Longwater, Green Park,
	Reading RG2 6GB, UK
	UK
Email:	mshand@cisco.com

Full Copyright Statement

[TOC](#)

Copyright © The IETF Trust (2008).

This document is subject to the rights, licenses and restrictions contained in BCP 78, and except as set forth therein, the authors retain all their rights.

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Intellectual Property

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the

procedures with respect to rights in RFC documents can be found in BCP 78 and BCP 79.

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.